4 Lecture 4

4.1 Operations on linear transformations

Definition 4.1 (Composition of linear transformations)**.** Let $T : V \to W$ and $S : W \to U$ be linear transformations. The **composition** ST (or $S \circ T$) of T and S is defined as

$$(ST)(x) = S(T(x)).$$

One can also check that $ST : V \to U$ is a linear transformation.

Proposition 4.2. *Let R, S, T be linear transformations. Whenever the composition in each item is defined, we have*
(i) $R(ST) = (RS)T$
(ii) $R(S + T) = RS + RT$
(iii) $(R + S)T = RT + ST$
(iv) For $a \in F$, $a(ST) = (aS)T = S(aT)$
(iv) $T \circ \mathrm{id}_V = T$, $\mathrm{id}_W \circ T = T$

Proof. One can directly verify these properties following the definition. □

Remark 4.3. We note that in general $ST \neq TS$ since as maps $T : V \to W$ and $S : W \to U$ one can only be composed in the order ST but not in the order of TS. Even when $V = W = U$, we do have examples of $T, S : V \to V$ such that $ST \neq TS$. $T = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $S = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Then $ST = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ and $TS = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.

4.2 Invertibility and rank-nullity theorem

Definition 4.4. Let $T : V \to W$ be a linear transformation.
T is **injective** (**one-to-one**) if $T(x) = T(y)$ implies $x = y$.
T is **surjective** (**onto**) if for any $y \in W$ there is $x \in V$ such that $T(x) = y$
T is **invertible** (**nonsingular**) if it is both injective and surjective. An invertible T is called an **isomorphism** between V and W.
V and W are **isomorphic** if there exists an isomorphism between them. We denote V and W are isomorphic by $V \cong W$.

Definition 4.5. The **kernel** (**null space**) of T is

$$\ker T = \{x \in V : T(x) = 0\}.$$

The **nullity** of T is $\operatorname{null} T = \dim \ker T$.
The **image** (**range**) of T is

$$\operatorname{im} T = \{T(x) : x \in V\}.$$

The **rank** of T is $\operatorname{rank} T = \dim \operatorname{im} T$.

Lemma 4.6. *Let $T : V \to W$ be a linear transformation. Then* $\ker T$ *is a subspace of* V *and* $\operatorname{im} T$ *is a subspace of* W.

Proof. Since $T(0_V) = 0_W$, $0_V \in \ker T$. Let $x, y \in \ker T$, $c \in F$. Then $T(x + y) = T(x) + T(y) = 0 + 0 = 0$ and $T(cx) = cT(x) = 0$. Thus $x + y \in \ker T$, $cx \in \ker T$.

Since $T(0_V) = 0_W$, $0_W \in \operatorname{im} T$. Let $u, v \in \operatorname{im} T$. Then there exist $x, y \in V$ such that $T(x) = u$ and $T(y) = v$. We have $u + v = T(x) + T(y) = T(x + y) \in \operatorname{im} T$ and $cu = cT(x) = T(cx) \in \operatorname{im} T$. □

Proposition 4.7. *We have the following characterization of injectivity and surjectivity of a linear transformation T.*

(i) T is injective if and only if $\ker T = \{0\}$.

(ii) T is surjective if and only if $\operatorname{im} T = W$.

Proof. (i) Since T is injective and $T(0) = 0$, we have $\ker T = \{0\}$.

Suppose $T(x) = T(y)$. Then by linearity of T, $T(x - y) = T(x) - T(y) = 0$. Thus $x - y \in \ker T = \{0\}$. Hence $x = y$.

(ii) This is just the definition of surjective. □

Example 4.8. $T : F^n \to F^m$, $T(x_1, \dots, x_n)^t = (x_1, \dots, x_r, 0, \dots, 0)^t$ for some $0 \le r \le \min(m, n)$. We have $\ker T = \{(0, \cdots, 0, x_{r+1}, \dots, x_n)^t : x_i \in F,\ r + 1 \le i \le n\}$ and $\operatorname{im} T = \{(x_1, \dots, x_r, 0, \dots, 0)^t : x_i \in F,\ 1 \le i \le r\}$. We have $\operatorname{null} T = n - r$ and $\operatorname{rank} T = r$. T is injective if $r = n$. T is surjective if $r = m$.

Proposition 4.9. *The following are equivalent:*

(i) $T : V \to W$ is invertible.

(ii) There is $S : W \to V$ linear such that $ST = \mathrm{id}_V$ and $TS = \mathrm{id}_W$.

(iii) T maps a basis into a basis.

If any of the above items is satisfied, then the linear transformation S in (ii) is unique.

Proof. (i)$\Longrightarrow$ (ii) Since T is both injective and surjective, we can define $S(y)$ to be the unique x such that $T(x) = y$. It follows from definition that $ST = \mathrm{id}_V$ and $TS = \mathrm{id}_W$.

We now show that S is linear. In $T(x + y) = T(x) + T(y)$, we take $x = S(u)$ and $y = S(v)$. Then we have $T(S(u) + S(v)) = TS(u) + T(S(v)) = u + v$. Composing S on both sides from the left, we get $S(u) + S(v) = S(u + v)$ Similarly we have $T(cS(u)) = cTS(u) = cu$ and composing S on both sides, we have $S(cu) = cS(u)$.

(ii)$\Longrightarrow$ (iii) Let $v_1, \dots, v_n$ be a basis of V. We show that $T(v_1), \dots, T(v_n)$ is a basis of W. To show that $T(v_1), \dots, T(v_n)$ is linearly independent, let $a_1T(v_1) + \cdots + a_nT(v_n) = 0$. We compose S on both sides and use the linearity of S and $ST = \mathrm{id}_V$, we have $S(a_1T(v_1) + \cdots + a_nT(v_n)) = a_1v_1 + \cdots + a_nv_n = 0 = 0$. Since $v_1, \dots, v_n$ is a basis, we have $a_1 = \cdots = a_n = 0$.

Let $y \in W$. Since $S(y) \in V$, there exist $a_1, \dots, a_n \in F$ such that $S(y) = a_1v_1 + \cdots + a_nv_n$. Then $y = TS(y) = T(a_1v_1 + \cdots + a_nv_n) = a_1T(v_1) + \cdots + a_nT(v_n)$. Thus $W = \operatorname{span}(T(v_1), \dots, T(v_n))$.

(iii) $\implies$ (i) To show that T is injective suppose $T(x) = 0$. We write $x = a_1v_1 + \cdots + a_nv_n$. Then by linearity, $a_1T(v_1)+\cdots+a_nT(v_n) = T(x) = 0$. Since $T(v_1), \dots, T(v_n)$ is a basis, $a_1 = \cdots = a_n = 0$. To show that T is surjective, let $y \in W$. Since $T(v_1), \dots, T(v_n)$ is a basis of W, $y = a_1T(v_1) + \cdots + a_nT(v_n) = T(a_1v_1 + \cdots + a_nv_n) \in \operatorname{im} T$.

To show uniqueness of S, we assume S_1, S_2 are two maps satisfying (ii). We consider the composition S_1TS_2. On one hand, $(S_1T)S_2 = \mathrm{id}_V S_2 = S_2$. On the other hand $S_1(TS_2) = S_1\mathrm{id}_W = S_1$. Thus $S_1 = S_2$. □

Definition 4.10. The linear transformation S in (ii) is called the **inverse** of T and is denoted by T^{-1}.

Corollary 4.11. *If $V \cong W$, then $\dim V = \dim W$.*

Proof. If $V \cong W$, then there is an isomorphism $T : V \to W$. By Proposition 4.9 (iii), T maps a basis into a basis and hence $\dim V = \dim W$. □

Lemma 4.12. *If $T; V \to W$ and $S : W \to U$ are both invertible, then ST is invertible and $(ST)^{-1} = T^{-1}S^{-1}$.*

Proof. We have $(ST)(T^{-1}S^{-1}) = S(TT^{-1})S^{-1} = SS^{-1} = \mathrm{id}_U$ and $(T^{-1}S^{-1})(ST) = T^{-1}(S^{-1}S)T = T^{-1}T = \mathrm{id}_V$. The conclusion follows from Proposition 4.9. □

Theorem 4.13 (First isomorphism theorem)**.** *Let $T : V \to W$ be a linear transformation. We define $\bar{T} : V/\ker T \to \operatorname{im} T$ to be $\bar{T}(x+\ker T) = T(x)$. Then $\bar{T}$ is a well-defined linear transformation and is an isomorphism.*

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "\pi"'] & W \\ V/\ker T \arrow[r, dashed, "\bar{T}"'] & \operatorname{im} T \arrow[u, "i"'] \end{tikzcd}$$

Proof. We first show that $\bar{T}$ is well-defined. For any $x' \in x+\ker T$, we have $x-x' \in \ker T$. Then $T(x) = T(x') + T(x - x') = T(x')$.

$\bar{T}$ is linear follows from the linearity of T.

If $\bar{T}(x + \ker T) = T(x) = 0$ then $x \in \ker T$. Thus $x + \ker T = 0 + \ker T$. This shows $\bar{T}$ is injective.

For any $y \in \operatorname{im} T$, there exists $x \in V$ such that $T(x) = y$. Thus $\bar{T}(x + \ker T) = y$. This shows that $\bar{T}$ is surjective. Hence $\bar{T}$ is an isomorphism. □

Corollary 4.14 (Rank-nullity theorem)**.** *Let $T : V \to W$ be a linear transformation. Then*

$$\dim V = \operatorname{null} T + \operatorname{rank} T.$$

Proof. By Theorem 4.13, $\bar{T} : V/\ker T \to \operatorname{im} T$ is an isomorphism. By we have $\dim(V/\ker T) = \dim \operatorname{im} T$. By Theorem 3.16, we have $\dim V - \dim\ker T = \dim(V/\ker T) = \dim \operatorname{im} T$. □

Corollary 4.15 (Underdetermined system of linear equations)**.** *Let $T : V \to W$ be a linear transformation and* $\dim W < \dim V$*. Then there is* $0 \neq x \in V$ *such that* $T(x) = 0$*.*

Proof. $\operatorname{null} T = \dim V - \operatorname{rank} T \geq \dim V - \dim W > 0$. □

Corollary 4.16. *Let V, W be finite dimensional vector spaces such that* $\dim V = \dim W$ *and $T : V \to W$ be a linear transformation. Then the following are equivalent:*

(i) T is an isomorphism.
(ii) T injective.
(iii) T surjective.
(iv) There is $S : W \to V$ such that $ST = \mathrm{id}_V$.
(v) There is $S : W \to V$ such that $TS = \mathrm{id}_W$.

Proof. (i) $\Longrightarrow$ (ii): This follows directly from definition.

(ii) $\Longrightarrow$ (iii): By (ii) and Lemma 4.7, we have $\operatorname{null} T = 0$. By Corollary 4.14, we have $\operatorname{rank} T = \dim V = \dim W$. By Lemma 2.17, $\operatorname{im} T = V$.

(iii) $\Longrightarrow$ (i): By (iii) we have $\operatorname{rank} T = \dim W = \dim V$. By Corollary 4.14, we have $\operatorname{null} T = 0$ and hence $\ker T = \{0\}$. Thus T is an isomorphism.

(i) $\Longrightarrow$ (iv) and (i) $\Longrightarrow$ (v): This follows from Proposition 4.9.

(iv) $\Longrightarrow$ (ii): Let $x \in \ker T$. Then $x = ST(x) = S(0) = 0$. Thus $\ker T = \{0\}$.

(v) $\Longrightarrow$ (iii): Let $y \in W$. Then $y = TS(y) = T(S(y)) \in \operatorname{im} T$. Thus $\operatorname{im} T = W$. □

Remark 4.17. The infinite dimensional version of (ii) and (iii) in this Corollary is called Fredholm alternative. It is useful in solving PDEs cf. [1, Chapter 8].

4.3 Basis, coordinates and matrix representation

In Lecture 2, we saw that a finite dimensional vector space is built up by a basis. In this section, we will see that given a basis on a vector space, a linear transformation reduces to a matrix.

Lemma 4.18 (Basis and coordinates)**.** *Let $v_1, \ldots, v_n$ be a basis of V. Then for every $v \in V$, there exists unique $x_1, \ldots, x_n \in F$ such that $v = x_1 v_1 + \cdots + x_n v_n$.*

Proof. Since $v_1, \ldots, v_n$ spans V, there exist $x_1, \ldots, x_n \in F$ such that $x = x_1 v_1 + \cdots + x_n v_n$. To show uniqueness we assume that $x = y_1 v_1 + \cdots + y_n v_n$. Then we have $(x_1 - y_1)v_1 + \cdots + (x_n - y_n)v_n = 0$. Since $v_1, \ldots, v_n$ is linearly independent, $x_1 - y_1 = \cdots = x_n - y_n = 0$. Thus $x_i = y_i$ for all $1 \leq i \leq n$. □

Definition 4.19. Let $\beta = (v_1, \ldots, v_n)$ be a basis of V. We say that $(x_1, \ldots, x_n)^t$ is the **coordinates** of v under the basis and we write $[v]_\beta$ for $(x_1, \ldots, x_n)^t$.

Theorem 4.20. *Let V be a n-dimensional vector space over F. We define the map $\phi_\beta : V \to F^n$ by $\phi_\beta(v) = [v]_\beta$. Then ϕ_β is a linear isomorphism and $\phi_\beta^{-1}(x_1, \ldots, x_n)^t = x_1 v_1 + \cdots + x_n v_n$.*

Proof. We first show that ϕ_β is linear. Let $\phi_\beta(v) = (x_1, \dots, x_n)^t$ and $\phi_\beta(w) = (y_1, \dots, y_n)^t$. Then $v+w = (x_1+y_1)v_1+\cdots+(x_n+y_n)v_n$ and hence $\phi_\beta(v+w) = (x_1+y_1, \dots, x_n+y_n)^t = \phi_\beta(v) + \phi_\beta(w)$. Moreover, $cv = cx_1v_1 + \cdots + cx_nv_n$ then $\phi_\beta(cv) = c\phi_\beta(v)$.

Next we show that ϕ_β is injective. If $\phi_\beta(v) = \phi_\beta(w) = (x_1, \dots, x_n)^t$, then $v = x_1v_1 + \cdots + x_nv_n = w$. Hence ϕ_β is injective.

Finally we show that ϕ_β is surjective. For any $(x_1, \dots, x_n)^t \in F^n$ we take $v = x_1v_1 + \cdots + x_nv_n$. Then $\phi_\beta(x) = (x_1, \dots, x_n)^t$.

Thus ϕ_β is an isomorphism and $V \cong F^n$.

Since $\phi_\beta(v) = (x_1, \dots, x_n)^t$ implies that $v = x_1v_1+\cdots+x_nv_n$, we see that $\phi_\beta^{-1}(x_1, \dots, x_n)^t = x_1v_1 + \cdots + x_nv_n$. □

Lemma 4.21. *Let $T : V \to W$ be a linear transformations and $v_1, \dots, v_n$ be a basis of V. Then T is uniquely determined by its values at $T(v_1), \dots, T(v_n)$.*

Proof. Let $v \in V$. Since $v_1, \dots, v_n$ is a basis of V, we have $v = x_1v_1 + \cdots + x_nv_n$. Then $T(v) = T(x_1v_1 + \cdots + x_nv_n) = x_1T(v_1) + \cdots + x_nT(v_n)$. □

Theorem 4.22. *Let $T : V \to W$ be a linear transformation. Let $\beta = (v_1, \dots, v_n)$ be a basis of V and $\gamma = (w_1, \dots, w_m)$ be a basis of W. Then there exists a unique matrix $A \in M_{m\times n}(F)$ such that for any $x \in F^n$,*

$$\phi_\gamma T \phi_\beta^{-1}(x) = Ax,$$

where $\phi_\beta : V \to F^n$ and $\phi_\gamma : W \to F^m$ are the coordinates isomorphisms.

The j-th column of A is given by the coordinates of $T(v_j)$ under the basis γ. That is $A = ([T(v_1)]_\gamma, \dots, [T(v_n)]_\gamma)$.

Proof. Since $(w_1, \dots, w_m)$ is a basis, for each j there exist scalars $a_{1j}, \dots, a_{mj}$ such that

$$T(v_j) = \sum_{i=1}^m a_{ij}w_i. \tag{1}$$

Now for any $x = (x_1, \dots, x_n)^t \in F^n$,

$$\begin{aligned} T\phi_\beta^{-1}(x) =& T(x_1v_1 + \cdots + x_nv_n) = \sum_{j=1}^n x_jT(v_j) \\ =& \sum_{j=1}^n x_j \sum_{i=1}^m a_{ij}w_i = \sum_{i=1}^m \Big(\sum_{j=1}^n a_{ij}x_j\Big)w_i. \end{aligned}$$

Thus

$$\phi_\gamma T\phi_\beta^{-1}(x) = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix} = Ax.$$

We note that by (1), the j-th column of A is given by the coordinates of $T(v_j)$ under the basis γ. This finishes the proof. □

Definition 4.23. The matrix A is called the **matrix representation** of T under the bases β of V and γ of W and is denoted by $[T]_\beta^\gamma$.

Corollary 4.24. *$\mathcal{L}(V,W) \cong M_{m\times n}(F)$. In particular,* $\dim \mathcal{L}(V,W) = (\dim V)(\dim W)$

Example 4.25 (The zero matrix)**.** For the zero map $0 : V \to W$, its matrix representation is the **zero matrix** under any basis β γ is $[0]_\beta^\gamma = 0_{m\times n}$.

Example 4.26 (The identity matrix)**.** For the identity map $\mathrm{id} : V \to V$, its matrix under a basis β is the **identity matrix** $[\mathrm{id}]_\beta^\beta = I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = (\delta_{ij})$ where $n = \dim V$ and we used the **Kronecker delta** symbol $\delta_{ij} = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$.

Example 4.27 (Diagonal matrix)**.** If $T(v_i) = \lambda_i v_i$ for all i, then

$$[T]_\beta^\beta = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Example 4.28. Let $T : \mathcal{P}_n(\mathbb{R}) \to \mathcal{P}_n(\mathbb{R})$ $Tf = f'$ and $\beta = (1, x, \ldots, x^n)$ be the standard basis. To compute $[T]_\beta^\beta$, we find $T(x^k) = kx^{k-1}$. Thus $[T(x^k)]_\beta = (0, \ldots, 0, k, 0 \ldots, 0)^t$ where the k is at the $(k-1)$-th position. Then

$$[T]_\beta^\beta = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 2 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & n \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

References

[1] Gilbarg D, Trudinger NS. Elliptic partial differential equations of second order. Berlin: springer; 1977.